They were Robbed! Scoring by the Middlemost to Attenuate Biased Judging in Boxing

Stuart Baumann[1] and Carl Singleton[2] iD

Abstract

Boxing has long grappled with the problem of biased or "bad" judging. At its worst, this leads to "Robberies", where boxers are widely seen as being denied rightful victories. Such incidents risk alienating fans and athletes. To address this problem, we propose a minimalist adjustment to the scoring system: the winner would be decided from the round-by-round scores of the judges, rather than relying on the judges' overall bout scores. This approach, known as consensus scoring, is rooted in social choice theory and utilises majority rule alongside middlemost aggregation functions. We show that this scoring method creates a coordination problem for actively partisan judges and theoretically attenuates their influence on fight outcomes. Our analysis and simulations, using a stylised model of strategic judging behaviour, demonstrate the potential of consensus scoring to significantly decrease the likelihood of a single partisan judge from swaying the result of a closely contested bout.

Keywords

scoring rules, judgement bias, contests, pugilism, combat sports

JEL Classification Codes: D91, L83, Z20, Z28

[1] Independent Researcher
[2] Economics Division, Stirling Management School, University of Stirling, Stirling, UK

Corresponding Author:
Carl Singleton, Economics Division, Stirling Management School, University of Stirling, Stirling, FK9 4LA, Scotland, UK.
Email: carl.singleton@stir.ac.uk

Introduction

Boxing has a reputation for partisan and corrupt judging. At the amateur level, some decisions in Olympic gold medal bouts have attracted criticism and ridicule, becoming boxing folklore, such as Roy Jones Jr.'s defeat in the 1988 (Seoul) light heavyweight final to a South Korean fighter (Ashdown, 2012), and Joe Joyce's defeat in the 2016 (Rio de Janeiro) super heavyweight final (Ingle, 2021; Rumsby, 2021). In professional boxing, there is longstanding suspicion about the integrity of judges (e.g., US Senate, 2001). Recent perceived "robberies" include Haney Vs. Lomachenko (Wainwright, 2023) and the first two editions of Alvarez Vs. Golovkin (Reid, 2023).

The prevalence of probable judging bias in combat sports has also been documented in a growing literature of empirical academic papers (e.g., Holmes et al., 2024; Lee et al., 2002). This behaviour was also described vividly in the recent judge-led independent investigation McLaren (2022) report, which examined unethical conduct in Olympic boxing after being commissioned by the Association Internationale de Boxe Amateur (AIBA). While the report did propose improved appointment processes and training of judges, it did not explore how to make the incentives inherent in the judging process more resilient to biases and corruption.

This short paper models the decisions of boxing judges and proposes an alternative scoring method that has the potential to significantly attenuate judge bias. Currently, scoring at the elite level is on a per-judge basis, with three judges usually employed for elite professional bouts and five at the Olympic amateur level. Under this system, a judge individually and subjectively scores each round according to the "10-Points Must" rules and then, in most cases, their entire "vote" goes to the boxer with the highest total score over all rounds. In most cases, the judge's vote then goes to the boxer they scored as winning a majority of rounds.[1] After the scorecards of the judges are collected, victory in a bout is then awarded to the boxer who receives the votes from a majority of judges. If neither boxer receives a majority of the judges' votes, due to at least one tied scorecard among the judges, then the bout is a draw. In this system, "aggregation over rounds and then judges", or "majority judges rule", it is relatively straightforward for a judge to ensure their vote goes to their favoured boxer. They just need to award them half the rounds (i.e., 7 of 12 for a world championship level men's professional bout). They can do this while minimising backlash, by choosing the best rounds for their favoured boxer.[2]

The change to the scoring system that we propose, "aggregation over judges and then over rounds", or "Consensus Scoring", is for each round to be awarded based on the aggregate scores over all judges. Normally, this would lead to whoever wins the majority of rounds winning the bout, rather than whoever wins on a majority of the judges' scorecards. This represents a minimalist change to the scoring system in the sport, so that the aggregation of judges' scores is first between them within rounds, and then over rounds, rather than vice versa. The minor nature of this change to the scoring system is sufficient to introduce a significant coordination problem for an actively partisan judge, and may be acceptable among fans.[3]

In fact, this rule change has been considered before in the sport of boxing. It was proposed in 2000 by the National Association of Attorneys General Boxing Task Force in the United States, which was set up after the Professional Boxing Safety Act became law in 1997 (National Association of Attorneys General Boxing Task Force, 2000). Motivated by an especially contentious decision in the heavyweight world championship bout between Evander Holyfield and Lennox Lewis in 1999, the Task Force suggested Consensus Scoring, based on a recommendation from a mathematician at Stanford University, Dr. Ralph S. Levine. Subsequently, Algranati and Cork (2000) evaluated the potential effects of changing to the Consensus Scoring rule, using the judges' scorecards from every professional world title bout between 1986 and 1999. They found that applying this scoring rule would have had minimal impacts on the outcome of these bouts. More recently, Berthet (2024) carried out a similar exercise with the same results by applying a consensus scoring rule on historical judging scorecards for thousands of mixed martial arts bouts, which have a similar scoring system to boxing.

Importantly though, while both of these ex post empirical studies recognised the potential benefits of Consensus Scoring in reducing the influence of anomalous judge scorecards on outcomes, neither could account for how a biased or partisan judge may have actually changed their scoring when faced with a Consensus Scoring system. We show that Consensus Scoring introduces a coordination problem for biased judges, which decreases their incentive to incorrectly award rounds towards a favoured boxer. This implies an inaccuracy in the counterfactual exercise of scoring bouts by consensus scoring when the observational data was generated by judges who were scoring under the majority judges system. We address this theoretically and with simulations, to demonstrate that the minimalist change to the status quo scoring system implied by Consensus Scoring could substantially alter the incentives and strategic behaviour of a partisan judge.

We focus on modelling the simplest practical case, with three judges for a bout, one of whom is biased in favour of one boxer. We also simplify our analysis by considering a hypothetical close fight, without knockdowns or point deductions, and where the rounds are all tight.[4] In this case, under majority judges rule, the presence of a partisan or biased judge can substantially increase the probability of a boxer winning despite being outnumbered by unbiased judges. Under Consensus Scoring, even if the partisan judge awards a majority of rounds to a favoured boxer, then this will have no impact on the final outcome unless those rounds align with the decisions of the other judges. This trivially mitigates the marginal influence of a single passively biased judge (e.g., a judge who has the same nationality as one of the fighters and displays nationality bias), whose anomalous scores become less relevant for the outcome of the bout. But more notably, the coordination problem implied by our proposed rule change also means that, to achieve a high probability of victory for their favoured boxer, the partisan judge, would have to decide to award more rounds to their favoured boxer than in the current system. This exposes them to scrutiny and potential backlash, as boxing pundits and fans will often criticise

poorly awarded rounds on judges' scorecards.[5] Our analysis and simulations of the model demonstrate that the scoring rule change could be highly effective in diminishing the incentives for biased judging in boxing and its influence on the outcomes of bouts.

From a theoretical perspective, the proposed Consensus Scoring rule is an application of majority rule and the middlemost aggregation function from social choice theory, which minimise the effective manipulability of outcomes by graders (e.g., Arrow, 1963; Balinski & Laraki, 2007; Young, 1974a, 1974b). This principle is already applied somewhat to the scoring in boxing, since the decision of the middlemost judge now determines the bout result. In Consensus Scoring, we instead suggest awarding bouts based on the aggregated middlemost round-by-round votes instead.

Our goal of improving the incentives of judges to score fights fairly in boxing is closely related to the focus of Frederiksen and Machol (1988), who analysed the subjective judging in sports like figure skating and dance, where the judges need to decide between multiple competitors, a setting where Arrow's (1963) theorem implies that all possible ways to combine judge preferences have some undesirable characteristics. Frederiksen & Machol proposed a new method for aggregating judge scores for such situations that attenuates some of these issues. Their context though faced the problem of the Arrow Impossibility Theorem (social choice paradox), given there were more than two alternative outcomes in the contest. That theorem does not apply here for a boxing bout since it consists of just two competitors, only one winner, and potentially biased judges.

In general, we contribute to the vast literature that either carries out post hoc analysis of changes to scoring rules and laws in sports or proposes new changes based on theory (for recent surveys see Kendall & Lenten, 2017; Wright, 2014). Our work falls into the latter type of study, particularly where minimalist changes have been proposed that could still in theory substantially improve the fairness of sports outcomes. For instance, in the world's most popular sport, association football, recent contributions have used simulations to explore whether incentives and outcomes could be altered significantly under different tie-breaking rules in round-robin tournaments (Csató, 2023; Csató et al., 2024), whether dynamic sequences in penalty shootouts could be fairer (Csató & Petróczy, 2022), and whether the allocation system for the additional slots of the expanded FIFA World Cup could be improved according to the stated goals of the organisers (Krumer & Moreno-Ternero, 2023).

Finally this paper builds on a growing literature studying various incentive issues in boxing and other combat sports (Akin et al., 2023; Amegashie & Kutsoati, 2005; Butler et al., 2023; Butler, 2023; Dietl et al., 2010; Duggan & Levitt, 2002; Tenorio, 2000). However, to the best of our knowledge, the incentives of boxing judges have not yet been studied directly, given the scoring rules they face, despite a well-developed literature on the influences and implications of biased decision making by the referees and judges in other sports (e.g., Bryson et al., 2021; Chowdhury et al., 2024; Dohmen & Sauermann, 2016; Reade et al., 2022, including other combat contests Brunello & Yamamura, 2023).

The remainder of our short paper proceeds as follows. In Section "The Model - A Partisan Judge in a Boxing Bout", we setup a stylised model of potentially biased judging and strategic behaviour in a boxing contest. Section "Analysis, Results, and Discussion" describes our analysis and discussion of the model. The detailed proofs of the main propositions regarding the scoring rules are presented in the Online Appendix, as are variations on the main results from simulating the model.

The Model - A Partisan Judge in a Boxing Bout

Consider a contest between two boxers of equal ability, in the Blue and Red corners. We assume each sequential round $t \in \{1, 2, , N\}$ of the contest has a true result, $\tau_t \in \{B, R\}$, which is a binomial random variable with equal probability.

Each judge, $j \in \{1, 2, 3\}$, gets an i.i.d. signal, $x_{t,j} \in \{B, R\}$, about the result of a round. With probability $\alpha \in (0, \frac{1}{2})$ this signal is the incorrect result, $x_{t,j} \neq \tau_t$, while with probability $1 - \alpha$ it is correct, $x_{t,j} \equiv \tau_t$.

Judges have a utility of:

$$U = S\mathbf{1}_{Blue\ wins} + G\mathbf{1}_{Red\ wins} - L, \quad L = \frac{\sum_{t=1}^{N} \mathbf{1}_{s_{t,j} \neq \tau_t}}{N} , \tag{1}$$

where $S \geq 0$ and $G \geq 0$ represents a judge's value from Blue or Red winning respectively. L refers to an expected backlash cost from biased judging.

We consider the case of two fair judges who have $S = G = 0$. As these judge's utility does not depend on the bout's winner, their optimal behaviour is to minimise backlash by awarding **fairly**, defined by choosing a round score of $s_{t,j} = B \Leftrightarrow x_{t,j} \equiv B$.[6] The third judge is actively or consciously **partisan** in favour of Blue and so has $S > 0$ and $G = 0$.

Under majority judges rule, the middlemost judge scorecard determines the bout. Under Consensus Scoring, the middlemost judge determines each round's winner, and then the middlemost round determines the bout. Judges award rounds separately and simultaneously.

Analysis, Results, and Discussion

The partisan judge ($j = 1$) can minimise backlash by awarding rounds fairly.[7] Under majority judges rule, they can maximally increase the chance of Blue winning the bout, while minimising backlash, by awarding $s_{t,1} = B$ in more than $\frac{N}{2}$ rounds. Under Consensus Scoring, their problem is more complex; a judge could award more than $\frac{N}{2}$ rounds to a boxer who then does not win them because the other judges disagreed.

If S is low, however, then the expected backlash can be sufficient for the partisan judge to award rounds fairly. We can characterise the critical $\hat{S}$ where the partisan

judge is indifferent between awarding fairly or gifting an additional round to *Blue*. We find that this critical value is higher under Consensus Scoring than majority judges rule, indicating that the former is more resilient to judge bias.

Proposition 1. For three-round bouts, in which *R*ed won a majority of rounds according to the true realisations, $\tau = [\tau_1, \tau_2, \tau_3]$, the critical $\hat{S}$ is higher under Consensus Scoring than majority judges rule, $\forall \alpha \in (0, \frac{1}{2})$.

Sketch of Proof: We can calculate the probability of each fair judge awarding a round for *Blue*, denoted by q, conditional on the signal seen by the partisan judge:

$$q|(x_{t,1} \equiv B) = (1 - \alpha)^2 + \alpha^2, \quad q|(x_{t,1} \equiv R) = 2\alpha(1 - \alpha). \tag{2}$$

Under Consensus Scoring, the number of fair judges awarding for *Blue* in a particular round can be represented as drawing from a binomial distribution with probabilities as in Equation 2. The survival function of this binomial, in conjunction with the decision of the partisan judge, is sufficient to infer the probability of *Blue* winning the round. From the probabilities for each round, we can derive the optimal number of rounds for the partisan judge to award for *Blue*. Under majority judges rule, we can infer the probability of another scorecard being in favour of *Blue* by combining the probabilities in Equation 2 across rounds. We can use these probabilities to evaluate whether the partisan judge should award additional rounds such that *Blue* wins on their card. Then we can derive, for each scoring rule, expressions for the critical $\hat{S}$ below which the partisan judge will award rounds fairly (see Online Appendix A). We find that there is a higher $\hat{S}$ under Consensus Scoring, giving us the proposition. ☐

Proposition 1 establishes that Consensus Scoring is more robust to partisan judging than the majority judges rule for three-round bouts. We numerically solve the model to establish the robustness of this result in longer bouts.[8] We use a benchmark parametrisation of $\alpha = 0.1$, $S = 0.8$, three judges (one of whom is partisan), and $N = 12$ rounds.[9]

To demonstrate a partisan judge's decision making, Figure 1 shows the probability of *Blue* winning the bout, given they truly won 6 rounds, for each number of rounds the partisan judge awards them. Under majority judges rule, there is a sharp increase in the probability of *Blue* winning if the partisan judge awards them more than 6 rounds. If *Blue* truly deserved to win 4 or 5 rounds, then, to award *Blue* the win, the partisan judge only needs to risk the backlash associated with giving them 3 or 2 more rounds on their scorecard. In contrast, Figure 1 shows that under Consensus Scoring, a judge cannot secure a sharp increase in the probability of *Blue* winning by giving them a small number of extra rounds; more rounds only gradually increase *Blue*'s chances.
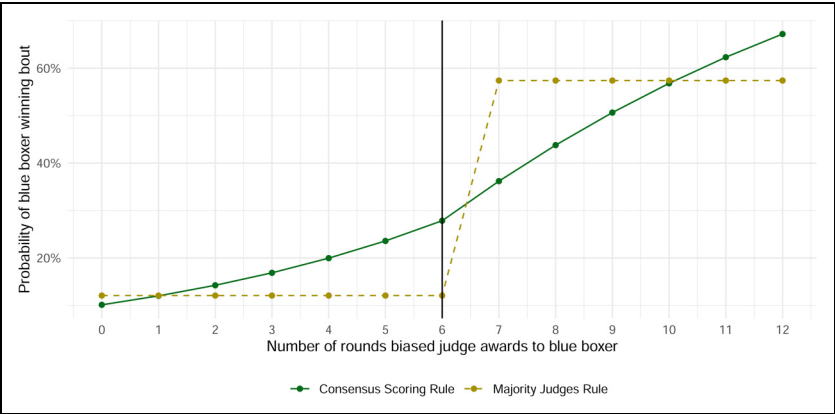
Figure 1. Simulated Probability of Blue winning, when both boxers truly won 6 of the 12 rounds, and 1 of the 3 judges favours Blue.
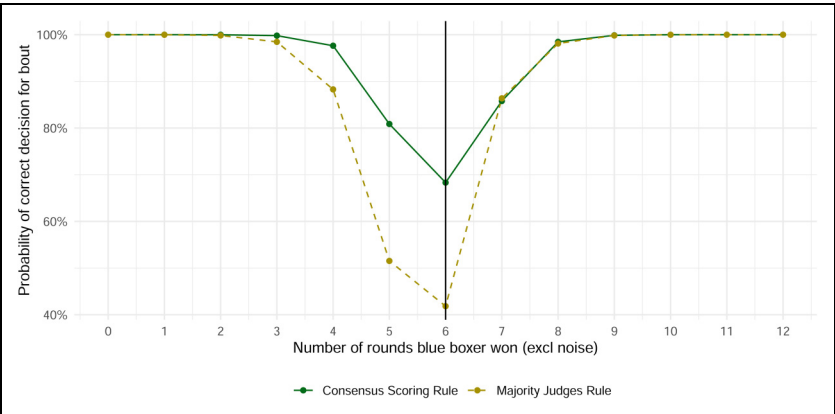

Figure 2. Probability of a "correct" result depending on the number of rounds truly won by Blue and how judges' scores are aggregated.

Figure 2 shows the impact of these differing incentives for the partisan judge, from running a series of simulations and counting the proportion of times each boxer wins under the two scoring systems, conditional on the true number of rounds won by Blue. When deciding the contest by majority judges, there is a high probability of erroneous results when Blue truly won only 4-6 rounds. When Blue truly wins the most rounds, the partisan judge unduly helps to lock in a deserved victory, so there is not a large difference in the number of incorrectly awarded bouts.

Finally, in the Consensus Scoring case, it can be noted from Figure 2 that the probability of the *B*lue boxer winning always increases when the biased judge awards them more rounds. This is in contrast to the majority judges case, when a judge ceases to impact the result at the point at which they award a majority of their card to a boxer. For instance, consider a bout where the fair judge sees 10 rounds with $x_{t,1} \equiv B$ and only two rounds are seen to be won by red. In this case, the biased judge may award additional rounds to *B*lue to lock in a *B*lue victory, while they would have no such incentive under majority judges rule.

This effect, however, does not tend to lead to a greater probability of an erroneous result under Consensus Scoring. The main reason for this is that the effect occurs in a context where *B*lue has likely won a large majority of rounds and is likely to win the bout. The more important case is when a bout is more even and there is a sharp increase, under the majority judges rule, in the winning probability at the 7 round level in Figure 2.

This point can be seen in Figure 3, which shows the probability of each possible outcome on the y-axis and the number of rounds *B*lue truly won (excluding noise) on the x-axis. Under majority judges rule (bottom panel), in evenly matched bouts, where the true result is a draw, *B*lue wins 47.0% and *R*ed wins 11.2%. When evenly matched bouts are awarded under Consensus Scoring (top panel), *B*lue wins 19.3% and *R*ed wins 13.4%.
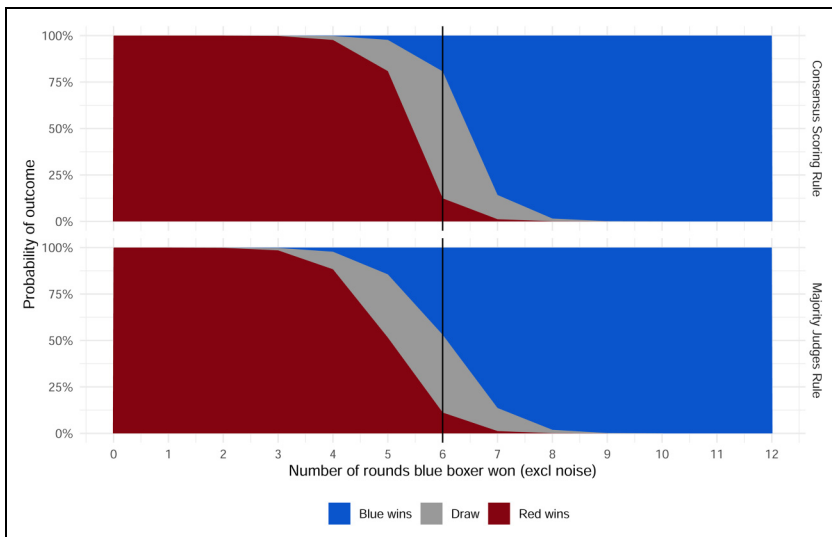

Figure 3. Probability of each outcome depending on the number of rounds truly won by *B*lue and how judges' scores are aggregated.

Figure 3 also shows the frequencies where one boxer wins despite the other deserving outright victory, e.g., the blue area to the left of the vertical black line. Under majority judges rule, it is more likely for an erroneous victory to be in favour of *B*lue than *R*ed; in this parametrisation, a robbery in favour of *B*lue is 12.5 times more likely than a robbery in favour of *R*ed. Under Consensus Scoring, the likelihood of a robbery is still in *B*lue's favour, by a multiple of 1.99, because there is still some incentive for the partisan judge to favour *B*lue. But this scoring system can substantially attenuate *B*lue's advantage from the presence of a partisan judge. There are also fewer robberies in absolute terms.

For robustness, the Online Appendices demonstrate extensions and checks on our analysis. Appendix C considers simulations with alternative parametrisations of the benchmark model, and, in Appendices D-F we repeat the analysis for setups consistent with women's professional, men's Olympic, and women's Olympic boxing, respectively (i.e., different numbers of rounds and judges). The results of all these extensions support our key findings: deciding bouts by Consensus Scoring, compared with by majority judges, makes it less likely that a partisan judge sways the outcome of a bout.

Acknowledgments

Declaration of Conflicting Interests

Funding

ORCID iD

Carl Singleton https://orcid.org/0000-0001-8247-8830

Supplemental Material

Supplemental materials for this article are available online.

Notes

1. It is possible for a boxer to earn a judge's vote without being scored as winning a majority of rounds, because the judge can score individual rounds as a draw or with a wider margin of victory than normal, based on knockdowns, fouls, or a particularly dominant performance in a round by one fighter.
2. The judge could also score some rounds with a wider margin of victory for their favoured boxer to ensure their victory even without awarding them a majority of rounds. However,

this could arguably attract more scrutiny since subjectively judged 10-8 rounds, that do not feature an objective knockdown, are rare.

3. Fans tend to scrutinise, oppose and criticise even quite small changes to the rules of their beloved sports. A notable example from cricket is the LBW rule, which has been continually 'improved' over the last century, often under opposition and criticism (Kumar, 2022).

4. Larger point difference rounds can be supported in the Consensus Scoring system using the same logic as in our simplified case - a round is scored as a 10-8 round only if the middlemost judges scores it as such. Though it would complicate our analysis here to allow judges more degrees of freedom in how they score rounds, the intuition of how the scoring system changes a biased judge's incentives and behaviour are not altered; A biased judge awarding an anomalous dominant round victory to one fighter would be mitigated by the coordination problem just the same.

5. Criticism of judges in social and sports media is generally fiercest after bouts where a robbery is perceived to have happened. It often focuses on specific rounds where a judge's decision appears to be particularly poor. While authorities seldom intervene to order a rematch, judges may be stripped of their status and not employed again (e.g., Slavin, 2017).

6. We analyse the case where all three judges are fair in Online Appendix B.

7. If $S \equiv 0$, then this judge's actions will be congruent to the fair judges.

8. This is done along the lines discussed in the sketch proof for Proposition 1. The code for calculating partisan judge best responses and bout simulations are included in the online supplementary material.

9. Importantly, Online Appendix B shows that when all three judges are fair, Consensus Scoring is still more accurate than the majority judges rule in generating a deserving winner of the bout. Intuitively this occurs as there are more combinations of rounds that could be flipped to change the result. Consider a three-round bout with three fair judges and a τ realisation of [**B**, **B**, **R**]. Consider that two mistakes happened in judging the bout (in that for two round-judges the $x_{t,j}$ realisation differs from τ). There are six possible pairs of $x_{t,j}$ values that can be flipped to change the result of the bout with Consensus Scoring. For one of the rounds where *B*lue won, we need to flip two $x_{t,j}$ values and there are six combinations that achieve this. But there are twelve possible pairs of $x_{t,j}$ values that can be flipped to change the result of the bout with the majority judges rule. We need to flip two of the $x_{t,j}$ values awarded to *B*lue on two different scorecards and there are twelve pairs of values that achieve this.

References

Akin, Z., Issabayev, M., & Rizvanoghlu, I. (2023). Incentives and strategic behavior of professional boxers. *Journal of Sports Economics*, *24*(1), 28–49.

Algranati, D. J., & Cork, D. L. (2000). Should professional boxing change its scoring system? A comparison of current and proposed methods. Carnegie Mellon H. John Heinz III School of Public Policy & Management 2000–06.

Amegashie, J. A., & Kutsoati, E. (2005). Rematches in boxing and other sporting events. *Journal of Sports Economics*, *6*(4), 401–411.

Arrow, K. J. (1963). *Social choice and individual values*. Yale University Press.

Ashdown, J. (2012). 50 stunning Olympic moments No14: Roy Jones Jr cheated out of gold. *The Guardian*. https://bit.ly/3SQ5KhL

Balinski, M., & Laraki, R. (2007). A theory of measuring, electing, and ranking. *Proceedings of the National Academy of Sciences*, 104(21), 8720–8725.

Berthet, V. (2024). Improving MMA judging with consensus scoring: A Statistical analysis of MMA bouts from 2003 to 2023. *arXiv preprint arXiv:2401.03280*.

Brunello, G., & Yamamura, E. (2023). Desperately Seeking a Japanese Yokozuna. Institute of Labor Economics (IZA) IZA Discussion Papers 16536.

Bryson, A., Dolton, P., James Reade, J., Schreyer, D., & Singleton, C. (2021). Causal effects of an absent crowd on performances and refereeing decisions during Covid-19. *Economics Letters*, 198(C), 109664.

Butler, R. (2023). An introduction to the james quirk special issue and the economics of combat sport. *Journal of Sports Economics*, 26(2), 135–138.

Butler, D., Butler, R., Maxcy, J., & Woodworth, S. (2023). Outcome uncertainty and viewer demand for basic cable boxing. *Journal of Sports Economics*, 26(2), 196–213.

Chowdhury, S. M., Jewell, S., & Singleton, C. (2024). Can awareness reduce (and reverse) identity-driven bias in judgement? Evidence from international cricket. *Journal of Economic Behavior & Organization*, 226: 106697.

Csató, L. (2023). How to avoid uncompetitive games? The importance of tie-breaking rules. *European Journal of Operational Research*, 307(3), 1260–1269.

Csató, L., Molontay, R., & Pintér, J. (2024). Tournament schedules and incentives in a double round-robin tournament with four teams. *International Transactions in Operational Research*, 31(3), 1486–1514.

Csató, L., & Petróczy, D. G. (2022). Fairness in penalty shootouts: Is it worth using dynamic sequences? *Journal of Sports Sciences*, 40(12), 1392–1398.

Dietl, H. M., Lang, M., & Werner, S. (2010). Corruption in professional sumo: An update on the study of duggan and levitt. *Journal of Sports Economics*, 11(4), 383–396.

Dohmen, T., & Sauermann, J. (2016). Referee bias. *Journal of Economic Surveys*, 30(4), 679–695.

Duggan, M., & Levitt, S. D. (2002). Winning isn't everything: Corruption in sumo wrestling. *American Economic Review*, 92(5), 1594–1605.

Frederiksen, J., & Machol, R. (1988). Reduction of paradoxes in subjectively judged competitions. *European Journal of Operational Research*. https://www.sciencedirect.com/science/article/abs/pii/037722178890375X

Holmes, B., McHale, I., & Zychaluk, K. (2024). Detecting individual preferences and erroneous verdicts in mixed martial arts judging using Bayesian hierarchical models. *European Journal of Operational Research*, 312(2), 733–745.

Ingle, S. (2021). Judges 'used signals' to fix Olympic boxing bouts, McLaren report finds. *The Guardian*. https://bit.ly/3sGXk1M

Kendall, G., & Lenten, L. J. A. (2017). When sports rules go awry. *European Journal of Operational Research*, 257(2), 377–394.

Krumer, A., & Moreno-Ternero, J. D. (2023). The allocation of additional slots for the FIFA world cup. *Journal of Sports Economics*, 24(7), 831–850.

Kumar, M. (2022). Explained: LBW rules and the controversial umpire's call in DRS. *The Times of India*. https://bit.ly/47F6ef4

Lee, H., Cork, D., & Algranati, D. (2002). Did lennox lewis beat evander holyfield? Methods for analysing small sample interrater agreement problems. *Journal of the Royal Statistical Society. Series D (The Statistician)*, *51*(2), 129–146.

McLaren, R. (2022). Independent investigation of the AIBA. https://bit.ly/3sHR5e1

National Association of Attorneys General Boxing Task Force (2000). Report. https://ag.ny.gov/sites/default/files/reports/report.pdf

Reade, J. J., Schreyer, D., & Singleton, C. (2022). Eliminating supportive crowds reduces referee bias. *Economic Inquiry*, *60*(3), 1416–1436.

Reid, A. (2023). Boxing's biggest robberies. *talkSPORT*. https://bit.ly/47BO6T8

Rumsby, B. (2021). Joe Joyce demands Rio 2016 gold medal from IOC after boxing corruption report. *The Telegraph*. https://bit.ly/46ndxqw

Slavin, H. (2017). Gennady Golovkin and Canelo Alvarez judge Adalaide Byrd disciplined for lopsided scorecard as she is stood down from major title fights. *Mail Online*. https://bit.ly/46nC8vh

Tenorio, R. (2000). The economics of professional boxing contracts. *Journal of Sports Economics*, *1*(4), 363–384.

US Senate (2001). Committee on commerce, science and transportation - A review of the professional boxing industry - is further reform needed? Senate Hearing 107-1090..

Wainwright, A. (2023). Fair or foul? Experts weigh in on Devin Haney-Vasiliy Lomachenko result. https://bit.ly/49IQUzO

Wright, M. (2014). OR analysis of sporting rules – A survey. *European Journal of Operational Research*, *232*(1), 1–8.

Young, H. (1974a). A note on preference aggregation. *Econometrica*, *42*(6), 1129–1131.

Young, H. (1974b). An axiomatization of Borda's rule. *Journal of Economic Theory*, *9*(1), 43–52.

Author Biographies

Stuart Baumann is currently an independent researcher, who received his PhD in Economics from the University of Edinburgh in 2017. His research interests include public economics, industrial organisation, macroeconomics, financial economics, consumer search, and sports.

Carl Singleton is currently Senior Lecturer of Economics at the University of Stirling. He is also an IZA Research Fellow. His research mainly focuses on macroeconomics and labour economics, with particular interests in business cycles, wage determination, and inequality issues. He has also published several papers about sports or using sports to test economic theory.